

AD-A239 638



MENTATION PAGE

Form Approved
OMB No 0704-0188

2

is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and reviewing the collection of information, Send comments regarding this burden estimate or any other aspect of this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Pike, Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

REPORT DATE

3. REPORT TYPE AND DATES COVERED

Final 1May1988 - 31May1991

4. TITLE AND SUBTITLE

Visual display of Quantitative Phenomena

5. FUNDING NUMBERS

DAAL03-88-K-0045

6. AUTHOR(S)

John W. Tukey

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

DTIC

SELECTE

AUG 20 1991

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

AR0 26098-23-MA-SDI

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Important aspects of the many sensor problem have been explored.
Shape theory has been consolidated and extended; new techniques have been studied.
Simple resistant smoothing in two (or more) dimensions has been studied and applied.
A graphical approach to exploratory analysis of variance, emphasizing single degrees of freedom has been developed and discussed.
Techniques for showing distributions of large samples and large letters in strips have been developed that are more sensitive and more flexible than previous techniques.
The importance of impact in designing graphical displays has been stressed; the likely evolution of graphical techniques has been described.

14. SUBJECT TERMS

Shape theory, shape distributions, landmarks, 2-D smoothing, head banging, graphical exploratory analysis of variance, distribution strips, limited randomization, designing for impact, future of graphical presentation

15. NUMBER OF PAGES

11

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT

UNCLASSIFIED

20. LIMITATION OF ABSTRACT

UL

**Best
Available
Copy**

Visual Display of Quantitative Phenomena

by

John W. Tukey

FINAL REPORT

May 1, 1988 — May 31, 1991

DAAL03-88-K-0045

supported by the

U. S. Army Research Office (Durham)

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or special
A-1	



91-08205



91 8 19 054

TABLE OF CONTENTS

Executive Summary	1
1. Personnel 1988 — 1991	2
2. Publications 1988 — 1991	2
Books, chapters in books	2
Papers	2
Submitted for publication:	3
Papers delivered and in preparation:	7
3. Technical Reports 1988 — 1991	5
Technical Reports: Department of Statistics Princeton University	5
Technical Reports: Department of Civil Engineering Princeton University	5
4. Shape Theory (Goodall)	6
5. Smoothing	6
6. Data analysis and display in the context of many sensors	7
7. Graphical methods in ANOVA	7
8. Graphic display of linked aspects	7
9. Columns divided into bars and their competitors	7
10. Designing graphics for impact	8
11. Visual display in the years to come	8
12. Graphical display of distributions: distribution strips	8
13. "Lo—tech" dynamic displays	8
14. Graphical techniques in multiple comparisons	8
15. Graphical techniques in describing plant-breeding trials	9
16. Mapping and other cartographic techniques	9

Executive Summary

Work under DAAL03-88-K-0045 focussed mainly on the following topics:

shape theory and techniques (5 items published, 2 to appear, 2 submitted, 4 delivered or in preparation);

smoothing (4 items published, 1 submitted);

graphical presentation of classes of data analytic results (1 item published, 1 submitted, 1 presented and to appear);

classical graphical techniques (4 technical reports);

distribution strips (1 item submitted, others in preparation);

visual display in the years to come (1 item published);

designing graphics for impact (1 item presented).

Special attention should be drawn to:

The Technical Report on summarization and iconization in the context of many sensors.

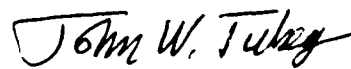
The work on consolidating and extending shape theory, especially as described in the RSS paper with discussion on procrustes methods, and the 4 papers "to appear" or "submitted".

The work on smoothing in two or more dimensions expanding the scope of "head banging" as a simple non-linear smoother.

The work on graphical methods in exploratory analysis of variance, which offer an important new approach focussed on single degrees of freedom, and later extended to the robust case (by Eugene Johnson, not supported by ARO).

The work on distribution strips, which is producing highly effective displays for large samples and, more generally for large batches.

The emphasis on impact as an essential element in the design of graphic displays.



John W. Tukey
Princeton 31 July 1991

1. Personnel 1988 — 1991

Faculty

Colin R. Goodall 1988-91

John W. Tukey 1988-91

Research Assistant

E. Olszewski 1988-91

2. Publications 1988 — 1991

Books, chapters in books:

Goodall, C. R. (1990). A survey of smoothing techniques. *Modern Methods in Data Analysis* (J. Fox and J. S. Long eds.). Sage, Beverly Hills, CA, pp. 126-176.

Goodall, C. R. and DeVeaux (1991). Final downsweeping: a use of Paull's rule for aggregation. *The Analysis of Variance, Vol. II*. (D. C. Hoaglin, F. Mosteller, and J. W. Tukey, eds.). John Wiley & Sons, Inc., New York. To appear.

Papers:

Goodall, C. R. (1988). Many New Jersey hospitals face loss of dollars under 1988 revised indirect cost reimbursement rules. *ECHO*, 34: 5-9.

Goodall, C. R. and Hansen, K. (1988). Resistant smoothing of high-dimensional data. *1988 Proceedings of the American Statistical Association: Statistical Graphics Section*. 17-21.

Goodall, C. R. and Lange, N. (1990). Growth curve models for correlated triangular shapes. *Proceedings of the Symposium on the Interface between Computer Science and Statistics*, Orlando, FL, pp 445-454.
[Originally titled: Analysis of shape differences in landmark data using multivariate growth curve models with random parameters]

Goodall, C. R. and Phelan, M. J. (1990). Edge-preservative smoothing and the assessment of point process models for GATE rainfall fields. *Statistical Inference in Stochastic Processes*. (N. U. Prabhu and I. V. Basawar, eds.) pp. 35-66.

Goodall, C. R. and Phelan, M. J. (1990). An assessment of a generalized WGR model for GATE rainfall. *J. Geophysical Research-Atmospheres*. 95: 7607-7615.

NOTE: Letters used with years on John Tukey's papers correspond to bibliographies in all volumes of his collected papers.

- Goodall, C. R. (1990). A simple objective method for determining a percent standard in mixed reimbursement systems. *J. Health Economics*, 9, 253-271.
- Goodall, C. R. (1990). One statistician's perspective. Discussion of D. Wahlsten Insensitivity of the analysis of variance to heredity-environment interaction. *Behavioral and Brain Sciences*, 13: 133-134.
- Goodall, C. R., Kim, S. H., Moss-Salentijn, L., and Shinozuka, M. (1990). Numerical simulation of the crown of an incisiform tooth by conformal and polynomial regression mapping of a simple model. *J. Theoretical Biology*. To appear.
- Goodall, C. R. and Studdiford, W. B. (1990). Comment on: Kaigh, W. D. (1989). A category representation paradox (with discussion). *American Statistician* 43: 92-97. *American Statistician* 44: 60-61.
- Goodall, C. R. (1991). Procrustes methods in the statistical analysis of shape (with discussion). *J. Royal Statistical Society, Series B*, 53: 285-339.
[originally titled: WLS estimators and tests for shape differences in landmark data.]
- Goodall, C. R. (1991). Eigenshape analysis of a cut-grow mapping for triangles, and its application to phyllotaxis in plants. *SIAM J. Applied Mathematics*. 51: (24 pp)
- Goodall, C. R. and Thoma, H. M. (1991). Graphic models and their description. *Proceedings of the Symposium on the Interface between Computer Science and Statistics*. (Seattle, Washington).
- Goodall, C. R. and Mardia, K. V. (1991). A geometrical derivation of the shape density. *Advances in Applied Probability*. To appear.
- Goodall, C. R. (1991). Graphical models in statistics. (A review of Whittaker, J. (1990) *Graphical Models in Applied Multivariate Analysis* Chichester, UK: John Wiley. To appear in *Technometrics* (5000 words).
- Goodall, C. R. and Mardia, K. V. (1991). The noncentral Bartlett decomposition and shape densities. *J. Multivariate Analysis*. To appear.
- Tukey, J. W. (1990). Data-based graphics: Visual display in the decades to come. *Statistical Science*. 5: 327-339.
- Tukey, J. W. (1991). The philosophy of multiple comparisons. *Statistical Science*. 6: 100-116.

Submitted for publication:

- Goodall, C. R. Lange, N., and Moss, M. L. (1991). Growth-curve models for repeated triangular shapes. Submitted to *Applied Statistics -- J. Roy. Stat. Soc. Series C*.
- Goodall, C. R. and Mardia, K. V. (1990). Multivariate aspects of shape theory. Submitted to *Annals of Statistics*.
- Kafadar, K., Goodall, C. R. and Tukey, J. W. (1991). An analysis of lung cancer mortality rates using urbanization and geography: Assessing sources of geographical variation. Revised, submitted to *International Statistical Review*.
- Johnson, E. G., and Tukey, J. W. (1991). Graphical methods for exploratory analysis of variance. Submitted to *Journ. Amer. Statist. Assoc.*
- Tukey, J. W. and Tukey, P. A. (1989). Dynamic graphic display of data can be "low tech" - - and projected. Submitted to *Amer. Statistician*.
- Tukey, P. A. and Tukey, J. W. (1990). Strips displaying empirical distribution: I. Textured dot strips. Submitted to *Journ. Amer. Statist. Assoc.*

Papers delivered and in preparation:

- Goodall, C. R. (1988). Techniques for geometric matching. Joint IMS-AMS-SIAM Summer Research Conference on Spatial Statistics and Imaging, Bowdoin College, ME.
- Goodall, C. R. and Tukey, J. W. (1988). Data analysis and iconization in the context of many sensors. I. Iconization, summarization, and available schemata. IST/SDIO Conference, Washington, D. C.
- Goodall, C. R. and Lange, N. (1989). Analysis of shape differences in landmark data using multivariate growth curve models with random parameters. Symposium on the Interface between Computer Science and Statistics.
- Goodall, C. R., DeVaux, R. D., and Phelan, M. J. (1989). Imaging geophysical phenomena on the SGI IRIS-4D. Northern New Jersey Chapter of the American Statistical Association.
- Goodall, C. R. and Winkelmann, J. (1989). The Princeton/HDR System Vendor presentation, Severity of Illness Symposium, State of New Jersey Department of Health, Princeton, NJ.
- Goodall, C. R. (1989). Regression models for directional data and their application to change in shape. Statistics, Earth, and Space Sciences, Leuven, Belgium.
- Goodall, C. R. (1989). Productive programming in S: ANOVA, smoothers, interpolation, and mapping. Inaugural S Conference, Wellington, New Zealand.
- Goodall, C. R. (1990). Procrustes methods and regression models for shape analysis. S. S. Wilks symposium on Shape theory and climate modelling. Princeton University, April.

- Tukey, J. W. (1988). Designing graphics for impact. Swiss Statistical Association, Bern, Switzerland, Universität Karlsruhe (TH), Federal Republic of Germany, and the Houston Area Chapter of the American Statistical Association (HACASA), Houston, Texas.
- Tukey, J. W. (1991). Where should multiple comparisons go next? Charles W. Dunnett Symposium, McMaster University, Hamilton, Ontario, Canada.

3. Technical Reports 1988—1991

Technical Reports: Department of Statistics, Princeton University

Number	Title	Author and date
295	Graphic comparisons of several linked aspects: alternatives and suggested principles	John W. Tukey October 1989 (in draft)
297	Columns divided into bars and their competitors	John W. Tukey July 1989 (in draft)
299	Scrawl strips, and letter of B-letter strips: depicting marginals of scatter plots (revised under DAAL03-88-K-0045)	John W. Tukey James G. Veitch May 1991
300	Combining CL (confidence limits) and PL (partial limits) about non-linear smoothers	John W. Tukey June 1990
301	Graphical displays of: Are the (x,y) pairs compatible with a linear dependence?	John W. Tukey June 1990

Technical Reports: Department of Civil Engineering, Princeton University

Title	Author and date
Data analysis and iconization in the context of many sensors. I. Iconization, summarization, and available schemata	Colin R. Goodall 1988

4. Shape Theory (Goodall)

Among the topics treated were:

Eigenshape analysis of an iterative mapping scheme applied to shapes of triangle

(paper in *SIAM J. Appl. Math.* 51)

Growth-curves models for the shapes of triangles of landmark points (joint with K. Lange and Moss, paper submitted to *Applied Statistics*)

Procrustes models for shapes, including an overview of shape theory and the connection of shape theory with the QR decomposition (paper with discussion, *J. Roy. Statist. Soc. B* 53:).

Probability models for shapes and forms; shape density (joint work with K. Mardia, papers (a) to appear in *Advances in Applied Probability*, (b) submitted to the *Annals of Statistics*, (c) submitted to *J. Multivariate Analysis*).

Conformal mapping of tooth shapes (joint with S. H. Kim, L. Moss-Salentijn, and M. Shinozuka, paper to appear in *J. Theoretical Biology*).

Exploring shape spaces with high-interaction graphics (paper at International Statistical Institute meeting, to appear in *Proceedings*).

Graphical models and their representation (to appear in *Proceedings of the 1991 Interface*).

5. Smoothing

A survey of smoothing techniques was prepared and published (Goodall).

A paper on resistant smoothing of irregularly spaced data by "head banging" was prepared and published (Goodall and Hansen).

A paper on head-banging techniques as applied to rainfall fields was reported and published (Goodall and Phelan).

A related paper on non-linear least squares and model assessment was prepared and published (Goodall and Phelan).

Smoothing and related topics were discussed at the inaugural S workshop in New Zealand (Goodall).

Investigation of smoothing county cancer rates continued. (Goodall, Kafadar, and Tukey).

6. Data analysis and display in the context of many sensors

A paper was written (Goodall and Tukey) and presented (at an IST/SDIO symposium, by Goodall). The new techniques were implemented in new S.

7. Graphical methods in ANOVA

A substantial paper on graphical and robust analysis of factorial data, making use of E. G. Johnson's new insights, has been written and submitted. (Johnson and Tukey).

8. Graphic display of linked aspects

A substantial technical report (#295) is in near final draft. (Tukey)

9. Columns divided into bars and their competitors

A substantial technical report (#297) is in near final draft. (Tukey)

10. Designing graphics for impact

A presentation has been drafted and used.

11. Visual display in the years to come

An invited paper for the Sesquicentennial of the American Statistical Association has appeared in their proceedings and in *Statistical Science*. (Tukey)

12. Graphical display of distributions: distribution strips

A technical report, on scrawl-strips and letter or B--letter strips, has been revised and issued (#299). (Tukey and Veitch)

A first paper on strip displays with a rectangular reference has been completed and submitted. (Tukey and Tukey)

A second paper, on strip displays with other references, is in a late stage of drafting. (Other papers are contemplated.)

13. "Lo--tech" dynamic displays

A paper has been completed and submitted

14. Graphical techniques in multiple comparisons

Work in this area forms part of two papers (The philosophy of multiple comparisons - - Where should multiple comparisons go next?) much of whose thrust is no explicitly graphical, but in which the importance of graphical output is stressed. (Tukey)

15. Graphical techniques in describing plant--breeding trials

Work with Kaye Basford (University of Queensland, Australia) has made substantial progress.

16. Mapping and other cartographic techniques

Extended correspondence between Mark Monmonier (Syracuse) and Tukey is raising some new issues and helping to settle old ones.